Tail bounds for gaps between eigenvalues of sparse random matrices[*]

Patrick Lopatto[†]   Kyle Luh[‡]

Abstract

We prove the first eigenvalue repulsion bound for sparse random matrices. As a consequence, we show that these matrices have simple spectrum, improving the range of sparsity and error probability from work of the second author and Vu. We also show that for sparse Erdős–Rényi graphs, weak and strong nodal domains are the same, answering a question of Dekel, Lee, and Linial.

Contents

1 Introduction

The gaps between eigenvalues of symmetric random matrices have been extensively studied by mathematicians and physicists. For the classical integrable ensembles, the Gaussian Orthogonal Ensemble and Gaussian Unitary Ensemble, the limiting spectral distribution follows the semicircle law. For an individual eigenvalue gap, however, the limiting distribution was only recently obtained [60]. Rapid progress in random matrix theory has permitted the extension of this result to a large class of random matrix models [2, 3, 11, 16, 17, 18, 19, 25, 26, 27, 28, 30, 37, 51, 57, 58, 59, 61, 69].

Much effort has been expended on understanding the extremal eigenvalue gaps, in particular the largest eigenvalue gap in the bulk of the spectrum, $\delta_{\max}$. Ben Arous and Bourgade [12] demonstrated that for the $n \times n$ GUE normalized so that its spectrum is supported on $[-2, 2]$, so that the typical inter-particle distance in the bulk is about n^{-1}, the largest bulk gap is of order $n^{-1}\sqrt{\log n}$. Figalli and Guionnet extended this result to β-ensembles with $\beta = 2$ [34]. In [32], Feng and Wei showed that the fluctuations of the largest gap are of order $n^{-1}\sqrt{\log n}$ and computed the limiting distribution. In work of the first author with Landon and Marcinek, the largest gap results of [12, 32] were extended to generalized Wigner matrices [38], including those with discrete entry distributions. We note that recent work of Bourgade [14], which presents a concise analysis of the convergence to equilibrium of Dyson Brownian motion, is able to recover the same result at the cost of imposing a weak smoothness assumption on the matrix entries.

While we now have a substantial understanding of the largest eigenvalue gap, the smallest gap, $\delta_{\min}$, is more difficult to investigate because it lies well below the typical inter-particle distance. Bourgade and Ben Arous [12] showed using the determinantal structure of the GUE that its smallest gap is of order $n^{-4/3}$. In [31], Feng, Tian, and Wei identified the normalized limit of the smallest eigenvalue gap of the GOE and found that the gap is of order $n^{-3/2}$; their argument builds on techniques previously developed by Feng and Wei to study circular β-ensembles [33]. Currently, the smallest gap lies outside of the purview of traditional universality results such as the Four Moment Theorem [62], and the techniques in the recent work [38] are not applicable. The strongest available result is in the recent work of Bourgade [14], which shows universality of the smallest gap, but requires that the matrix entries possess a weak form of smoothness. At present, no universality results exist for the smallest gap for matrices that are sparse or have discrete entry distributions, such as a matrix of Bernoulli random variables.

While tail bounds are known for the individual gaps when the matrix entries are more general random variables [58, 61], the error rates are not strong enough to take a union bound to conclude anything about the minimum gap. We now scale the matrices so that their spectrum lies on $[-2\sqrt{n}, 2\sqrt{n}]$, which makes the average inter-particle distance $n^{-1/2}$; we take this convention to match the existing tail bound literature, and it remains in force throughout the rest of the paper. For Hermitian matrices, under stringent smoothness and decay assumptions on the random variables, a result of Erdős, Schlein, and Yau [29] implies that there exists a small constant $c > 0$ such that

$$\mathbb{P}\left(\delta_{\min} \leq \frac{\delta}{n^{1/2}}\right) = o(n\delta^3) + \exp(-cn)$$

for any $\delta > 0$. For discrete random variables, it was a milestone just to show that $\delta_{\min} > 0$ [63]. In particular, Tao and Vu showed that for any $A > 0$, with probability at least $1 - n^{-A}$ a random symmetric matrix has simple spectrum, meaning every eigenvalue appears with multiplicity one. In follow-up work with Nguyen [48], they showed the following tail bound for the eigenvalue gaps. Given eigenvalues λ_i labeled in ascending order, we denote the gaps by $\delta_i = \lambda_{i+1} - \lambda_i$.

Theorem 1.1 ([48, Theorem 2.1]). *There exists a constant $c > 0$ such that the following holds for the eigenvalue gaps, δ_i, of a real symmetric Wigner matrix. For any $n^{-c} \leq \alpha \leq c$ and $\delta \geq n^{-c/\alpha}$,*

$$\sup_{1 \leq i \leq n-1} \mathbb{P}\left(\delta_i \leq \frac{\delta}{n^{1/2}}\right) = O\left(\frac{\delta}{\alpha^{1/2}}\right).$$

Setting $\alpha = n^{-c}$, one can deduce that a real symmetric random matrix has simple spectrum with probability at least $1 - O(\exp(-n^c))$. A related problem, posed by Babai, is whether the adjacency matrix of an Erdős–Rényi random graph has simple spectrum. This was resolved affirmatively for all dense random graphs in [48, 63]. A consequence in complexity theory is that for such random graphs the graph isomorphism problem is in complexity class $\mathcal{P}$ [6].

In this work we study the eigenvalue gaps of sparse random matrices. The theory of sparse random matrices is of interest in its own right, but it also has innumerable applications in computer science and statistics. In contexts where sparse random matrices have similar spectral guarantees as their dense counterparts, they offer significant advantages as they require less space to store, allow quicker multiplication, and need fewer random bits to generate [5, 7, 8, 21, 22, 47]. A popular model for such matrices is to consider the Hadamard (entrywise) product of a dense random matrix and a sparse matrix of independent (up to symmetry) indicator variables with expectation $p = p(n)$. Much work has been done to transfer the results known for dense random matrices to the sparse setting [9, 10, 15, 37, 39, 42, 55, 68]. Although the results resemble their dense analogues, the sparsity brings about a variety of complications in the proofs. Only recently, the second author and Vu showed that for a large class of random variables and for $p \geq n^{-1+\varepsilon}$ with $\varepsilon > 0$, a sparse random matrix has simple spectrum with probability at least $1 - O_\varepsilon(\exp(-(np)^{1/128}))$ [43], where this notation indicates that the implied constant depends on ε. This implies that the graph isomorphism problem restricted to this class of sparse random graphs is in complexity class $\mathcal{P}$.

Our main contribution is to go beyond verifying such matrices have simple spectrum and prove a tail bound for the minimal eigenvalue gap of sparse random matrices with $p \geq C \log^{7+\varepsilon}(n)/n$. In comparison with [43], our results represent an improvement in both error probability and the range of sparsity considered. As an application of our tail bound, we show that for sparse Erdős–Rényi graphs, weak and strong nodal domains are the same, answering a question of Dekel, Lee, and Linial [24]. Our results also expand the range of sparse graphs for which the graph isomorphism problem is known to be in $\mathcal{P}$. Related to this last application is the graph matching problem, for which various algorithms contingent on simple spectrum are known [1, 44, 65]; our results similarly extend their range of applicability.

2 Main Results

We begin with a formal definition of our random matrix model.

Definition 2.1. *We let M_n denote a symmetric random matrix with entries*

$$m_{ij} = \xi_{ij}\chi_{ij},$$

where the ξ_{ij} are independent (for $i \geq j$), mean zero, variance one, and subgaussian with subgaussian moment B, and the χ_{ij} are independent (for $i \geq j$) Bernoulli random variables with $\mathbb{E}\chi_{ij} = p$.

Theorem 2.2. *Let M_n be as in Definition 2.1, and fix $\nu > 0$. There exist constants $C_{2.2}, c_{2.2}, c'_{2.2} > 0$, depending only on the subgaussian moment B, such that for*

$$\frac{C_{2.2} \log^{7+\nu} n}{n} \leq p \leq 1$$

and

$$(np)^{-1/(7+\nu)} \leq \alpha \leq \frac{c'_{2.2}}{\log n},$$

the following holds for the gaps between the eigenvalues, $\delta_i = \lambda_{i+1} - \lambda_i$. For any $\delta \geq \exp(-\alpha^{-1})$,

$$\sup_{1 \leq i \leq n-1} \mathbb{P}\left(\delta_i \leq \delta \exp\left(-c_{2.2}\frac{\log(1/p)}{\log np}\right)\sqrt{\frac{p}{n}}\right) \leq C_{2.2}\frac{\delta}{\alpha}.$$

Observe that there is a trade-off in the strength of the error bound and the size of the eigenvalue gap, determined by the value of α. For example, if we choose $\alpha = c_{2.2}/\log n$, we obtain the following result.

Corollary 2.3. *Let M_n be as in Definition 2.1, and fix $\nu > 0$. There exist $C_{2.3}, C'_{2.3} > 1$, such that for $p \geq \frac{C_{2.2} \log^{7+\nu} n}{n}$,*

$$\sup_{1 \leq i \leq n-1} \mathbb{P}\left(\delta_i \leq \delta \exp\left(-c_{2.2}\frac{\log(1/p)}{\log np}\right)\sqrt{\frac{p}{n}}\right) \leq C_{2.3}\delta \log n.$$

for $\delta \geq n^{-C'_{2.3}}$. By a union bound,

$$\mathbb{P}\left(\delta_{\min} \leq \frac{\sqrt{p}}{n^{3/2+o(1)}}\right) = o(1).$$

At the other extreme, setting $\alpha = (np)^{-1/(7+\nu)}$ and $\delta = \exp(-\alpha^{-1})$, we have the following result.

Corollary 2.4. *Let M_n be as in Definition 2.1, and fix $\nu > 0$. For $p \geq \frac{C_{2.2} \log^{7+\nu} n}{n}$,*

$$\mathbb{P}(M_n \text{ has eigenvalues with multiplicity}) \leq \exp\left(-\frac{1}{2}(np)^{1/(7+\nu)}\right).$$

Observe that when $p = 1$, which is the dense case considered in [48], the above two corollaries recover [48, Corollary 2.2] and [48, Corollary 2.3], which are the analogous extreme cases of the bound in [48, Theorem 2.1].

Remark 2.5. This result improves the range of sparsity in [43] from $n^{-1+\varepsilon}$ for some $\varepsilon > 0$ to $\log n^{7+\nu}/n$. Even in the regime $p \geq n^{-1+\varepsilon}$, our result improves on the bound in [43] where the probability of not having a simple spectrum was less than $\exp(-(np)^{1/124})$. However, we suspect that the optimal bound should be $\exp(-cnp)$ for some constant $c > 0$. The sparsity range of Theorem 2.2 is near optimal as $p = o(\log n/n)$ yields multiple rows and columns entirely of zeros. This generates repeated eigenvalues at 0.

We also have the same result for adjacency matrices of random Erdős–Rényi graphs. Let $G(n,p)$ denote the random graph on n vertices with edges appearing independently and with probability p.

Theorem 2.6. *Let A_n be the adjacency matrix of the random Erdős–Rényi graph $G(n,p)$, and fix $\nu > 0$. There exist constants $C_{2.2}, c_{2.2}, c'_{2.2} > 0$, depending only on the subgaussian moment B, such that for*

$$\frac{C_{2.2} \log^{7+\nu} n}{n} \leq p \leq 1 - \frac{C_{2.2} \log^{7+\nu} n}{n}$$

and

$$(np)^{-1/(7+\nu)} \leq \alpha \leq \frac{c'_{2.2}}{\log n},$$

the following holds for the gaps between the eigenvalues, $\delta_i = \lambda_{i+1} - \lambda_i$. For any $\delta \geq \exp(-\alpha^{-1})$,

$$\sup_{1 \leq i \leq n-1} \mathbb{P}\left(\delta_i \leq \delta \exp\left(-c_{2.2} \frac{\log(1/p)}{\log np}\right) \sqrt{\frac{p}{n}}\right) \leq \frac{\delta}{\alpha}. \tag{2.1}$$

Remark 2.7. Note that an upper bound on p is necessary in this case as $p = 1$ generates a *deterministic* matrix with repeated eigenvalues. Additionally, our argument can be easily applied to random perturbations of a finite rank matrix; see Remark 6.2. However, for perturbations of an arbitrary matrix, new ideas are needed as many of the delicate net arguments cannot be adapted when the operator norm of the perturbed matrix is large. For dense random graphs, this was done in [48, Theorem 2.6].

2.1 Non-degeneration of eigenvectors and nodal domains of a random graph

Consider the eigenfunctions of the Laplacian on a Riemannian manifold. The zero sets of these eigenfunctions partition the space into so-called *nodal domains*. These domains are of great interest to geometers and have been intensively studied (see [20, 40, 46] and the references therein). Here we consider a discrete analogue, the nodal domains of eigenvectors for adjacency matrices of random graphs, which has its roots in graph theory and has recently found uses in data science [23, 24, 35]. Given an eigenvector u of an adjacency matrix A, we call a subset D of the vertices a *weak nodal domain* if it is connected, $u(x)u(y) \geq 0$ for $x, y \in D$, and D is a maximal subset under these two conditions. A *strong nodal domain* is defined similarly using the strict inequality $u(x)u(y) > 0$. Dekel, Lee, and Linial conjectured that the notions of strong and weak domains are equivalent for random graphs [24], and this was shown for $G(n,p)$ with constant p in [48]. A consequence of the following non-degeneration result is that we are able to resolve this conjecture for $p \geq C_{2.2} \log^{7+\nu}(n)/n$.

Theorem 2.8. *Let A_n be the adjacency matrix of the random graph $G(n,p)$, and fix $\nu > 0$. For any $D > 0$, there exists a $C = C(D) > 0$ such that for*

$$\frac{C_{2.2} \log^{7+\nu} n}{n} \leq p \leq 1 - \frac{C_{2.2} \log^{7+\nu} n}{n},$$

the probability that there exists an eigenvector $v = (v_1, \ldots, v_n)$ of A_n with $|v_i| \leq n^{-C}$ for some i is at most Cn^{-D}.

Theorem 2.8 provides a quantitative lower bound on the mass of the eigenvector components, complementing the vast literature on eigenvector delocalization, which provides upper bounds (see [50, Section 4] and [13]).

Corollary 2.9. *For any $D > 0$, there exists $C = C(D) > 0$ such that with probability at least $1 - Cn^{-D}$, the strong and weak nodal domains of $G(n,p)$ are the same.*

Arora and Bhaskara [4] showed that for random graphs $G(n, p)$ with $p \geq n^{-c}$, where c is a constant that may be determined explicitly,[1] all non-first eigenvectors of the adjacency matrix A_n of $G(n, p)$ have exactly two weak nodal domains with high probability. Recall that since the adjacency matrix is not centered, the eigenvector corresponding to the largest eigenvalue behaves differently, tending to align itself with the all ones vector [45]. Combining this result with our previous corollary yields the following simple statement.

Corollary 2.10. *There exists $c > 0$ such that the following holds. For any $D > 0$ and $p \geq n^{-c}$, there exists $C = C(D) > 0$ such that with probability at least $1 - Cn^{-D}$, each eigenvector of $G(n, p)$ (except the first) has exactly two strong nodal domains which partition the vertices.*

An identical non-degeneration result applies to matrices M_n defined in Definition 2.1.

Theorem 2.11. *Fix $\nu > 0$. For any $D > 0$, there exists a $C = C(D) > 0$ such that for*

$$p \geq \frac{C_{2.2} \log^{7+\nu} n}{n}$$

the probability that there exists an eigenvector $v = (v_1, \ldots, v_n)$ of M_n with $|v_i| \leq n^{-C}$ for some i is at most Cn^{-D}.

Remark 2.12. Theorems 2.8 and 2.11 represent specific examples of a range of possible results. Specifically, varying α in Theorem 2.2 can lead to trade-offs in the size of the entries and the strength of the probability bound. We have chosen to give a simple polynomial bound on the size and probability for the sake of simplifying the presentation.

We also remark that nodal domains were studied in the recent work [36], which showed that there exists a constant $c \geq 0$ such that for $p \geq n^{-c}$ the two nodal domains identified in [4] are balanced, meaning they each contain close to $n/2$ vertices with high probability. Further, [54] shows that, with high probability, any vertex is connected to some vertex in the other domain.

The remainder of the paper is organized as follows. In Section 3, we outline the key steps and intuition for the proof of Theorem 2.2. In Sections 4 and 5, we prove several preliminary results about eigenvectors of sparse random matrices. In Section 6.1, we provide the proof of Theorem 2.2. In Section 6.2 we provide the necessary modifications to extend Theorem 2.2 to non-centered random matrices, such as the adjacency matrices of Erdős–Rényi graphs, proving Theorem 2.6. Finally, in Section 6.3, we prove Theorem 2.8.

3 Proof Strategy

The proof follows the same broad outline as [43]. For M_n as in Definition 2.1, we decompose the matrix as

$$M_n = \begin{pmatrix} M_{n-1} & X \\ X^T & m_{nn} \end{pmatrix}, \tag{3.1}$$

where $X = [x_1, \ldots, x_{n-1}] \in \mathbb{R}^{1 \times (n-1)}$. For a matrix W, let $\lambda_n(W) \geq \cdots \geq \lambda_1(W)$ be the eigenvalues of W. Fix an integer i such that $1 \leq i \leq n$ and let $v = (x, a)$ (where $x \in \mathbb{R}^{n-1}$ and $a \in \mathbb{R}$) be the unit eigenvector associated to $\lambda_i(M_n)$. By definition we have

$$\begin{pmatrix} M_{n-1} & X \\ X^T & m_{nn} \end{pmatrix} \begin{pmatrix} x \\ a \end{pmatrix} = \lambda_i(M_n) \begin{pmatrix} x \\ a \end{pmatrix}.$$

---

[1]The authors give an exact value. However, the published version of an eigenvector delocalization estimate used to prove the result differs slightly from the version given in [4], where it is cited by the authors in pre-publication form. The value of the constant should be adjusted in light of this.

For the top $n-1$ coordinates this gives (writing $\lambda_i(M_n)$ for $\lambda_i(M_n)\,\mathrm{Id}$)

$$(M_{n-1} - \lambda_i(M_n))x + aX = 0.$$

Let w be the eigenvector of M_{n-1} corresponding to $\lambda_i(M_{n-1})$. Multiplying on the left by w^T, we obtain

$$|aw^T X| = |w^T(M_{n-1} - \lambda_i(M_n))x| = |\lambda_i(M_{n-1}) - \lambda_i(M_n)||w^T x|. \tag{3.2}$$

By the Cauchy interlacing theorem, we have $\lambda_i(M_n) \le \lambda_i(M_{n-1}) \le \lambda_{i-1}(M_n)$.

Since the entries of M_n are subgaussian, we have with high probability that

$$\lambda_i \in [-K\sqrt{pn}, K\sqrt{pn}]$$

for some constant K that depends only on the subgaussian moment B of the entries. Therefore, the average size of an eigenvalue gap is roughly $O\left(\frac{\sqrt{pn}}{n}\right) = O\left(\sqrt{\frac{p}{n}}\right)$. For any $\hat{\delta} > 0$, let $\mathcal{E}_i = \mathcal{E}_i\left(\hat{\delta}\right)$ denote the event that

$$\lambda_{i+1} - \lambda_i \le \hat{\delta}\sqrt{\frac{p}{n}}.$$

We also let $\mathcal{G}_i$ be the intersection of the event $\mathcal{E}_i$ with the event that the eigenvector $v = (x, a)$ with eigenvalue λ_i has $|a| \ge n^{-1/2}$. Therefore, by (3.2) and using $|w^T x| \le 1$, on the event $\mathcal{G}_i$, we have

$$|w^T X| \le \hat{\delta}\sqrt{p}. \tag{3.3}$$

We wish to show this is unlikely.

Recall that the theory of small ball probability (e.g. [49]) examines the probability that a random variable takes values in a small interval. Therefore, we have reduced the problem to understanding the small ball probability of the inner product of a random vector with the eigenvector w. It is known that this small ball probability is related to the amount of "disorder" in the coordinates of the eigenvector. Broadly speaking, a large amount of disorder implies the small ball probability is small. We deal with the case that w has high disorder eigenvectors using these results. To exclude all eigenvectors with low disorder, we employ a covering argument, varying our approach according to the structure of the eigenvector.

The covering argument is completed in multiple stages. For a fixed λ, we consider $M_n - \lambda\,\mathrm{Id}$ acting on the unit sphere, where Id is the identity operator. Following the prescription initiated in a series of works [9, 10, 41, 53, 56, 64], we decompose the sphere into several sets that each offer their own advantages. *Compressible* vectors are those vectors that are close to m-sparse vectors for some parameter m. In [9], it was shown that the product of the matrix with a compressible vector has many large coordinates and therefore large ℓ_2 norm. We adapt this argument to our symmetric matrix case to exclude compressible vectors. We next consider *dominated* vectors, which are those vectors whose coordinates outside the m largest coordinates have a small ratio of ℓ_2 norm to ℓ_∞ norm. This type of vector was introduced in [9]. As these vectors are also nearly sparse, they can be excluded similarly to the compressible vectors.

Finally, for vectors that are neither compressible nor dominated, we use a stratification according to a measure of structure, the LCD. The LCD was introduced in [56] and is defined later. As our random matrix is symmetric, there is dependence between the rows which prevents us from applying small ball probability estimates to each coordinate independently.[2] To address this problem, for a fixed v we partition the coordinates

---

[2]This obstacle is what prevents us from reaching the optimal threshold for p by simply following the argument in [9], which considered non-symmetric matrices for $p \ge (C\log n)/n$.

of v into small subsets; this is similar to the method used in [66]. For a fixed subset, after conditioning on the columns of $M_n - \lambda$ outside of the subset, we can extract more independent coordinates to use in small ball estimates. There is some flexibility in the size of these subsets, and this ultimately results in the trade-off between the error probability and gap size in Theorem 2.2.

The previous steps are done for a fixed λ and hold with exponentially high probability. Taking a union bound over a fine enough net of the interval $[-K\sqrt{pn}, K\sqrt{pn}]$ completes the argument.

A similar approach was applied in [43], under the assumption that $p \geq n^{-1+\varepsilon}$ for some $\varepsilon > 0$ and therefore small polynomial terms could often be neglected. In our current setting, where p is on the order of $\log^C n/n$, it turns out that the above decomposition is insufficient primarily because the vectors that are not dominated or compressible can have a wide range of ℓ_2 mass in their coordinates outside of the m largest. Therefore, we further decompose the vectors by their ℓ_2 mass in the relevant coordinates. Working in each of these classes allows some key technical estimates that bypass the small polynomial losses from [43]. These technical improvements generate the improvement in the range of sparsity and the error probability. Furthermore, in [43], the result was only concerned with a non-zero separation of the eigenvalues. A more careful accounting of the small ball probability greatly improves the (implicit) small ball estimate in [43].

4 Compressible and Dominated Vectors

The goal of this section is to prove Proposition 4.6, which shows that any eigenvector of M_n cannot be close to a sparse vector, in a certain quantitative sense (with high probability). Before proceeding to its proof, we introduce a few necessary definitions and lemmas.

4.1 Decomposition of the sphere

We now formally define the decomposition of the unit sphere used in the proof sketch of Section 3.

Definition 4.1. *Fix $m < n$. The set of m-sparse vectors is given by*

$$\mathrm{Sparse}(m) = \{x \in \mathbb{R}^n : |\operatorname{supp}(x)| \leq m\}.$$

Furthermore, for $\delta > 0$, we define the compressible and incompressible vectors by

$$\mathrm{Comp}(m, \delta) = \{x \in \mathbb{S}^{n-1} : \exists y \in \textit{Sparse}(m) \text{ such that } \|x - y\|_2 \leq \delta\},$$

and

$$\mathrm{Incomp}(m, \delta) = \{x \in \mathbb{S}^{n-1} : x \notin \mathrm{Comp}(m, \delta)\}.$$

For any $1 \leq n \leq n'$, we let $[n]$ denote the set $\{1, 2, \ldots, n\}$ and $[n : n']$ denote the set $\{n, n+1, \ldots, n'\}$.

Definition 4.2. *For any $x \in \mathbb{S}^{n-1}$, let $\pi_x : [n] \to [n]$ be a permutation which arranges the absolute values of the coordinates of x in non-increasing order. For $1 \leq m \leq m' \leq n$ denote by $x_{[m:m']} \in \mathbb{R}^n$ the vector with coordinates*

$$x_{[m:m']}(j) = x_j \cdot \mathbb{1}_{[m:m']}(\pi_x(j)).$$

For any $c < 1$ and $m \leq n$, define the set of vectors with dominated tail by

$$\mathrm{Dom}(m, c) = \{x \in \mathbb{S}^{n-1} : \|x_{[m+1:n]}\|_2 \leq c\sqrt{m}\|x_{[m+1:n]}\|_\infty\}.$$

This definition was first given in [9]. Like compressible vectors, vectors with dominated tail are close to being sparse, though in a different way. This approximate sparsity facilitates the proof of the following key bound, Proposition 4.4.

4.2 Bounds for compressible and dominated vectors

We first state a high probability bound on the operator norm of M_n, which was defined in Definition 2.1.

Lemma 4.3 ([43, Proposition 5.2] and [67, Proposition 1.10]). *For M_n defined in Definition 2.1, there exist constants $C_{4.3}, K, c_{4.3} > 0$, depending only on the subgaussian moment B, such that for $p \geq \frac{C_{4.3} \log n}{n}$ and $n \geq (c_{4.3})^{-1}$,*

$$\mathbb{P}(\|M_n\| \geq K\sqrt{pn}) \leq \exp(-c_{4.3} pn).$$

For the remainder of this work, all references to the constant K refer to the K provided by Lemma 4.3.

The compressible and dominated vectors were previously resolved in [43] down to the optimal scale $p \geq C \log n / n$. Given some $\bar{C}_{4.4} > 0$, we define the parameters

$$\ell_0 = \left\lceil \frac{\log 1/(8p)}{\log \sqrt{pn}} \right\rceil, \qquad \rho = (\bar{C}_{4.4})^{-\ell_0 - 6}.$$

Proposition 4.4 ([43, Proposition 5.3]). *There exist constants $C_{4.4}, \bar{C}_{4.4}, c_{4.4}, c'_{4.4}, > 0$, depending only on the subgaussian moment B of Definition 2.1, such that the following holds. If p, m, λ satisfy*

$$p \geq \frac{C_{4.4} \log n}{n}, \quad p^{-1} \leq m \leq c_{4.4} n, \text{ and } \lambda \in [-K\sqrt{pn}, K\sqrt{pn}], \tag{4.1}$$

then with probability at least $1 - \exp(-c'_{4.4} pn)$,

$$\|(M_n - \lambda)x\|_2 \geq c_{4.4} \rho \sqrt{pn}$$

for all $x \in \mathrm{Comp}(m, \rho) \cup \mathrm{Dom}(m, c'_{4.4})$ and $n > (c'_{4.4})^{-1}$.

Remark 4.5. Note that if $p \geq n^{-1+c}$ for some constant $c > 0$, then ρ is bounded below by a constant. At the optimal scale $p = C \log n / n$, there exist constants $C_1, C_2, c_1, c_2 > 0$ such that

$$C_1 \exp(-c_1 \log n / \log \log n) \leq \rho \leq C_2 \exp(-c_2 \log n / \log \log n).$$

We now come to the main result of this section, which combines the previous two proposition to exclude the possibility of compressible or dominated eigenvectors.

Proposition 4.6. *Let be M_n as in Definition 2.1 with $p \geq C_{4.4} \frac{\log n}{n}$. For $p^{-1} \leq m \leq c_{4.4} n$ and $n \geq (c_{4.6})^{-1}$,*

$$\mathbb{P}(\text{there exists an eigenvector } v \in \mathrm{Comp}(m, \rho) \cup \mathrm{Dom}(m, c_{4.4})) \leq \exp(-c_{4.6} pn)$$

for some constant $c_{4.6} > 0$.

Proof. Let $\mathcal{N}$ denote a $c_{4.4} \rho \sqrt{pn}$-net of the interval $[-K\sqrt{pn}, K\sqrt{pn}]$ with

$$|\mathcal{N}| \leq \frac{4K}{c_{4.4} \rho}. \tag{4.2}$$

If there exists a compressible or dominated eigenvector v with eigenvalue $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$, then there exists a $\lambda_0 \in \mathcal{N}$ such that

$$\|(M_n - \lambda_0)v\|_2 = \|(\lambda - \lambda_0)v\|_2 \leq c_{4.4} \rho \sqrt{pn}.$$

By a union bound and Proposition 4.4, the probability of this event is bounded by

$$|\mathcal{N}| \exp(-c_{4.4} pn) \leq \exp(-c_{4.6} pn).$$

for large enough $C_{4.4}$ and small enough $c_{4.6}$; to bound $|\mathcal{N}|$, we used Remark 4.5. Finally, the event that that there exists an eigenvalue outside of the interval $[-K\sqrt{pn}, K\sqrt{pn}]$ is bounded by $\exp(-c_{4.3} pn)$, by Lemma 4.3. Shrinking $c_{4.6}$ allows us to take a union bound to include this event, and concludes the proof. $\qquad \square$

5 Incompressible Vectors

In this section, we show that M_n does not have structured eigenvectors. We begin with Section 5.1, where we elucidate the connection between small ball probability and our measure of structure, the Least Common Denominator (LCD). Section 5.2 and Section 5.3 are devoted to the proof of Proposition 5.17, which shows it is unlikely an eigenvector of M_n has an LCD lying in a given level set. This proposition is the main technical achievement of this section. Finally, we derive Proposition 5.18 as a straightforward consequence of Proposition 5.17 and a union bound, which excludes the possibility of structured eigenvectors altogether. Together with Proposition 4.6, Proposition 5.18 will allow us to complete the outline of Section 3 and prove our main theorems in the next section.

5.1 Small ball probability

Recall from the proof sketch in Section 3 that we wish to bound the probability that the inner product of an eigenvector and a random vector is small. This motivates the definition of Lévy concentration, which bounds the small ball probabilities of a random vector Z.

Definition 5.1. *The* Lévy concentration *of a random vector $Z \in \mathbb{R}^n$ is defined to be*

$$\mathcal{L}(Z, \varepsilon) = \sup_{u \in \mathbb{R}^n} \mathbb{P}(\|Z - u\|_2 \leq \varepsilon).$$

When X is a random vector and v is a fixed vector, the structure of v will greatly influence the Lévy concentration of the random variable $v \cdot X$. To formalize this concept, we begin with a measure of arithmetic structure for a unit vector.

Definition 5.2 ([66, Definition 6.1]). *Let p be as in Theorem 2.2. We define the least common denominator (LCD) of $x \in \mathbb{S}^{n-1}$ as*

$$D(x) = \inf \left\{ \theta > 0 : \mathrm{dist}(\theta x, \mathbb{Z}^n) < \left(\frac{\log_+(\sqrt{\gamma p}\theta)}{\gamma p} \right)^{1/2} \right\},$$

where γ is an appropriate constant that is defined in Remark 5.3 below.

Remark 5.3. There exist constants $\gamma, \bar{\varepsilon}_0 \in (0, 1)$ such that for any $\varepsilon \leq \bar{\varepsilon}_0$,

$$\mathcal{L}(\xi\chi, \varepsilon) \leq 1 - \gamma p,$$

where χ is a Bernoulli random variable such that $\mathbb{P}(\chi = 1) = p$ and ξ is a subgaussian random variable with unit variance. We fix such a γ in Definition 5.2.

Proposition 5.4 ([9, Proposition 4.2]). *Let $X \in \mathbb{R}^n$ be a random vector with i.i.d. coordinates of the form $\xi_j \chi_j$, where the χ_j's are Bernoulli random variables with $\mathbb{P}(\chi_j = 1) = p$ and the ξ_j's are random variables with unit variance and finite fourth moment. Then for any $v \in \mathbb{S}^{n-1}$,*

$$\mathcal{L}\left(X \cdot v, \sqrt{p}\varepsilon\right) \leq C_{5.4}\left(\varepsilon + \frac{1}{\sqrt{p}D(v)} \right),$$

where $C_{5.4}$ depends only on the fourth moment of ξ.

We may tensorize Proposition 5.4 to obtain a bound on the Lévy concentration of $M_n x$. The argument is almost identical to the proof of [9, Proposition 4.3], and we note only the necessary modifications here. Recall the notation $x_{[m:m']}$ from Definition 4.2. For any index set $J \subset [n]$, we extend this notation to x_J in the canonical way.

Proposition 5.5 (Small ball probabilities of $M_n x$ via regularized LCD)**.** *There exists a constant $C_{5.5}$ such that for any $\alpha, \varepsilon > 0$ and index set I of size $\lceil \alpha n \rceil$,*

$$\mathcal{L}(M_n x, \varepsilon \|v_I\|_2 \sqrt{pn}) \leq C_{5.5}^{n-\lceil \alpha n \rceil} \left(\varepsilon + \frac{1}{\sqrt{p} D(v_I / \|v_I\|_2)} \right)^{n-\lceil \alpha n \rceil}.$$

Proof Sketch. We first observe that conditioning on elements of M_n never decreases (and may increase) $\mathcal{L}(M_n x, \varepsilon \|v_I\|_2 \sqrt{pn})$. We therefore condition on all elements not in columns indexed by elements of I, and also condition the elements whose indices (i, j) satisfy $i, j \in I$. The remaining elements are i.i.d. and consist of $n - \lceil \alpha n \rceil$ rows. The remainder of the argument is nearly identical to the one leading to [9, Proposition 4.3], where an analogous statement was shown for non-symmetric matrices. □

The following lemma provides a lower bound for the LCD in terms of the ℓ^∞ norm.

Proposition 5.6 (Lemma 6.2, [66])**.** *For all $x \in \mathbb{S}^{n-1}$,*

$$D(x) \geq \frac{1}{2\|x\|_\infty}.$$

As in [66], we define a regularized version of the LCD. However, our definition is slightly different than the one in [66]. Recall the notation $\text{Incomp}(m, \delta)$ given after Definition 4.1, and observe that the set I_0 in the following definition takes a distinguished role and is not included in the maximum. Here, k_0 represents a parameter that will be fixed later, in the material preceding (5.2).

Definition 5.7 (Regularized LCD)**.** *Let $\{I_j\}_{j=0}^{k_0}$ be any partition of $[n]$ with k_0 elements.. We define the* regularized LCD *of a vector $v \in \text{Incomp}(m, \delta)$ as*

$$\widehat{D}(v) = \widehat{D}(I, v) = \max_{1 \leq j \leq k_0} D\big(x_{I_j} / \|x_{I_j}\|_2\big).$$

In our use of Definition 5.7 below, I_0 will be (approximately) the m largest coordinates of v. Hence $\widehat{D}(v)$ gives a measure of the structure of the elements of v left over after approximating v by an m-sparse vector.

5.2 Decomposition of incompressible vectors

In this section, we define a way to decompose incompressible vectors, which is used in the proof of Proposition 5.15 below. In order to give this decomposition, we first introduction a classification of the incompressible vectors, which allows us to control the amount of mass that is not in the m largest coordinates.

Definition 5.8. *For $\rho \leq \rho_1 \leq \rho_2 \leq 1$ and $c < 1$, define*

$$\text{Incomp}_{\rho,c}(m, \rho_1, \rho_2) =$$
$$\left\{ v \in \mathbb{S}^{n-1} \cap (\text{Comp}(m, \rho) \cup \text{Dom}(m, c))^c : \rho_1 \leq \|v_{[m+1:n]}\|_2 < \rho_2 \right\}.$$

Remark 5.9. By definition, $\|v\|_2 \leq \rho$ for any $v \in \text{Comp}(m, \rho)$, which gives rise to the condition $\rho \leq \rho_1$ in the preceding definition.

We will consider the sets of incompressible vectors $\text{Incomp}_{\rho,c'_{4.4}}(m, 2^{j-1}\rho, 2^j\rho)$ for $j \in \mathbb{N}$, where m is a parameter that will be chosen later. For brevity, we introduce the shorthand

$$\text{Incomp}(m, 2^{j-1}\rho, 2^j\rho) = \text{Incomp}_{\rho,c'_{4.4}}(m, 2^{j-1}\rho, 2^j\rho).$$

For the remainder of this section we primarily use the fact that the vectors in $\text{Incomp}(m, 2^{j-1}\rho, 2^j\rho)$ are not dominated. That they are not compressible is used only in the proof of Proposition 5.17.

We begin with a straightforward upper bound. Recall ρ was defined in Proposition 4.4. Fix $j \in \mathbb{Z}$ and consider a vector $v \in \mathrm{Incomp}(m, 2^{j-1}\rho, 2^j\rho)$. Since $v \notin \mathrm{Dom}(m, c'_{4.4})$,

$$\|v_{[m+1:n]}\|_2 > c'_{4.4}\sqrt{m}\|v_{[m+1:n]}\|_\infty.$$

Furthermore, since $\|v_{[m+1:n]}\|_2 < 2^j\rho$ by definition,

$$c'^{-1}_{4.4}\frac{2^j\rho}{\sqrt{m}} > \|v_{[m+1:n]}\|_\infty. \tag{5.1}$$

On the other hand, we can also find a large set of coordinates that are uniformly lower-bounded.

Lemma 5.10. *For $v \in \mathrm{Incomp}(m, 2^{j-1}\rho, 2^j\rho)$, the set*

$$\sigma(v) = \left\{ i \in [n] : |v_i| \geq \frac{2^{j-1}\rho}{2\sqrt{n}} \text{ and } i \in \pi_v^{-1}([m+1:n]) \right\}$$

satisfies $|\sigma(v)| \geq (c'_{4.4})^2 m/8$.

Proof. For the sake of contradiction, assume that $|\sigma(v)| < (c'_{4.4})^2 m/8$. Then by (5.1),

$$\|v_{[m+1:n]}\|_2 \leq \sqrt{\|v_{[m+1:n]}\|_\infty^2 |\sigma(v)| + n\frac{2^{2(j-1)}\rho^2}{4n}} < 2^{j-1}\rho,$$

contradicting the definition of $\mathrm{Incomp}(m, 2^{j-1}\rho, 2^j\rho)$. $\qquad\square$

We now define a partitioning procedure. For this, we introduce some new notation.

Definition 5.11. *For a set $I \in [n]$ with $|I| \geq k_2 > k_1$, we use $I_{\langle k_1:k_2\rangle}$ to denote all the elements from the k_1-th to the k_2-th in I (inclusive), where we order the elements from least to greatest. For example, if $I = \{2, 4, 5, 6, 9\}$ then $I_{\langle 2:4\rangle} = \{4, 5, 6\}$.*

Let $v \in \mathbb{S}^{n-1}$ be a vector, let $\omega = \omega(n)$ be a parameter satisfying

$$n^{-1/7} \leq \omega \leq \frac{1}{\log n},$$

and set $m = \omega n$. We define k_0 as the largest number of disjoint subsets with $\lceil \omega n \rceil$ elements one can have of $[n]$ whose union does not contain the indices of the m largest elements of v. We consider disjoint index sets $I_1, \ldots, I_{k_0}$, each of size $\lceil \omega n \rceil$, each not containing any indices of the m largest elements of v. Therefore,

$$\frac{1}{2\omega} \leq \left\lfloor \frac{n-m}{\lceil \omega n \rceil} \right\rfloor = k_0 \leq \frac{1}{\omega}. \tag{5.2}$$

In our definition, the index sets I_j depend on v, but we suppress this dependence in the notation. For a vector $v \in \mathbb{S}^{n-1}$, let $\tau(v)$ denote the set of indices of the m largest coordinates. By Lemma 5.10, we can choose a subset $\widehat{\sigma}(v) \subset \sigma(v)$ of size exactly $\lceil (c'_{4.4})^2 m/8 \rceil$, where $\sigma(v)$ was defined in the statement of that lemma. We observe that $\widehat{\sigma}(v)$ and $\tau(v)$ are disjoint.

Let $\overline{\sigma}(v) = [n] \setminus (\tau(v) \cup \widehat{\sigma}(v))$. For $1 \leq k < k_0$, we define

$$r' = \left\lceil \frac{(c'_{4.4})^2 m}{8} \right\rceil, \quad r = \left\lfloor \frac{r'}{k_0} \right\rfloor, \quad s = \lceil \omega n \rceil - r,$$
$$I_k = \widehat{\sigma}(v)_{\langle 1+(k-1)r:kr\rangle} \cup \overline{\sigma}(v)_{\langle 1+(k-1)s:ks\rangle}. \tag{5.3}$$

For the rest of this work, we drop floor and ceiling functions because they do not influence the argument in a substantial way.

Finally, we define $I_0 = [n] \setminus \cup_{k=1}^{k_0} I_k$. In words, I_0 contains the m largest coordinates and the smaller coordinates left over from divisibility issues. In particular, $|I_0| \leq m + \lceil \omega n \rceil$. Since the sets I_k were chosen to be disjoint for $k \geq 1$, it follows that $\{I_k\}_{k=0}^{k_0}$ is a partition of $[n]$.

The primary objective of this partition is recorded in the following lemma, where we also define the constants ρ'_j.

Lemma 5.12. *For $v \in \mathrm{Incomp}(m, 2^{j-1}\rho, 2^j\rho)$ and $1 \leq k \leq k_0$,*

$$\rho'_j := c'_{4.4} 2^{j-3} \rho\omega \leq \|v_{I_k}\|_2 \leq c'^{-1}_{4.4} 2^j \rho = 2^3 (c'_{4.4})^{-2} \rho'_j \omega^{-1}. \tag{5.4}$$

Also,

$$\widehat{D}(v) \geq (c'_{4.4})^2 2^{-5} n^{1/2} \omega^{3/2}.$$

Proof. The bounds on $\|v_{I_k}\|_2$ follow from the coordinate-wise bounds of our construction. For the lower bound, we ignore all elements not in $\widehat{\sigma}(v)$. We obtain

$$\sqrt{\frac{|\widehat{\sigma}(v)|}{k_0}} \frac{2^{j-1}\rho}{2\sqrt{n}} \leq \|v_{I_k}\|_2 \leq \sqrt{|I_k|} \|v_{[m+1:n]}\|_\infty.$$

The claim (5.4) then follows from Lemma 5.10, (5.1), and (5.2).

For the second claim, applying Proposition 5.6 and recalling Definition 5.7 yields

$$\widehat{D}(v) \geq \min_{k \geq 1} \left\{ \frac{\|v_{I_k}\|_2}{2\|v_{I_k}\|_\infty} \right\}. \tag{5.5}$$

Then the claim follows from the lower bound on $\|v_{I_k}\|_2$ in the previous paragraph and (5.1). $\qquad\square$

5.3 Vectors with small LCD

We now exclude vectors with small regularized LCD as potential eigenvectors of M_n. This is the content of the next proposition, Proposition 5.15, which shows that any vector in $\mathrm{Incomp}(m, 2^{j-1}\rho, 2^j\rho)$ with small regularized LCD is unlikely to be near an eigenvector. We first define level sets of vectors according to their regularized LCD.

Definition 5.13. *For any $L > 0$, we define the level sets*

$$S_L = \{v \in \mathrm{Incomp}(m, \rho) : L \leq \widehat{D}(v) < 2L\}.$$

We also require a preliminary lemma. Recall γ was defined in Remark 5.3.

Lemma 5.14 (Lemma 6.13, [43]). *Let $\omega > 0$, and let $f(n)$ be a function such that*

$$\lim_{n \to \infty} f(n) = \infty.$$

Then for $L > f(n)$, the set of unit vectors

$$\{v \in \mathbb{S}^{\omega n - 1} : f(n) \leq D(v) \leq L\}$$

admits a β-net of size at most

$$\left(12 + \frac{\bar{c}L}{\sqrt{\omega n}}\right)^{\omega n} \log(L)$$

where $\bar{c} > 0$ is a universal constant and

$$\beta = \frac{2\sqrt{\log(2\sqrt{\gamma p}L)}}{L\sqrt{\gamma p}}.$$

We now state and prove the main technical result of this section. Recall S_L was defined in (5.13), ρ'_j was defined in Lemma 5.12, and K is the constant given by Lemma 4.3.

Proposition 5.15. *Fix $\nu > 0$. There exist constants $C_{5.15}, c_{5.15}, c'_{5.15}, \tilde{c}_{5.15} > 0$ such that for $p \geq C_{5.15} \frac{\log^{7+\nu} n}{n}$, $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$, $j \in \mathbb{N}$, and for any*

$$(np)^{-1/(7+\nu)} \leq \alpha \leq \frac{c'_{5.15}}{\log n}$$

and

$$\tilde{c}_{5.15} \alpha^{3/2} n^{1/2} \leq L \leq p^{-1/2} \exp(\alpha^{-1}),$$

the following holds for $n \geq (c'_{5.15})^{-1}$:

$$\mathbb{P}\Big(\exists v \in \mathrm{Incomp}(\alpha n, 2^{j-1}\rho, 2^j \rho) \cap S_L \text{ s.t. } \|(M_n - \lambda)v\|_2 \leq c_{5.15} \varepsilon_0 \rho'_j \sqrt{pn}\Big) \leq \exp(-c'_{5.15} n),$$

where

$$\varepsilon_0(L) = \min\left\{\frac{c'_{5.15}\sqrt{\alpha n}}{L}, \frac{c'_{5.15}\sqrt{\log r}(\log\log n)}{\alpha^2 r}\right\} \text{ and } r = \frac{\tilde{c}_{5.15}}{2}\alpha^{3/2}(np)^{1/2}.$$

Proof. We set $m = \alpha n$, and define

$$\mathcal{K} = \mathrm{Incomp}(m, 2^{j-1}\rho, 2^j \rho) \cap S_L.$$

In outline, this proof implements the following steps:

1. Construct a suitable net $\mathcal{M}$ for $\mathcal{K}$.
2. Upper bound the size of $\mathcal{M}$.
3. Show the claim holds for all $v \in \mathcal{M}$.
4. Extend the result from all $v \in \mathcal{M}$ to all $v \in \mathcal{K}$.

For Step 1, let $v \in \mathcal{K}$ be a vector and consider the partition $\{I_k\}_{k=0}^{k_0}$ of the coordinates of v constructed in (5.3) with the parameter $\omega = \alpha$. For the coordinates I_0, by a standard volume estimate,[3] there exists a $c'_{5.15}\rho'_j\varepsilon_0/10K$-net, $\mathcal{N}_0$, of the values $[0, 1]$ such that

$$|\mathcal{N}_0| \leq \left(\frac{30K}{c'_{5.15}\varepsilon_0\rho'_j}\right)^{m+\alpha n},$$

where we recall $|I_0| \leq m + \alpha n$.

For the coordinates in I_k with $k \geq 1$, we use a construction that exploits the LCD structure. Observe that the hypothesis of Lemma 5.14 holds for $v_{I_k}/\|v_{I_k}\|_2$ because

$$D(v_{I_k}/\|v_{I_k}\|_2) \geq (c'_{4.4})^2 \frac{m}{2^5}\sqrt{\frac{\alpha}{n}} = \frac{c'_{4.4}}{2^5}\alpha^{3/2}n^{1/2}, \tag{5.6}$$

as shown in the proof of Lemma 5.12 (see (5.5)), and the lower bound tends to infinity as $n \to \infty$. For I_k with $k \geq 1$, let $\mathcal{N}_k$ denote the β-net guaranteed by Lemma 5.14 applied to $v_{I_k}/\|v_{I_k}\|_2$.[4]

We next implement a net of scaling factors. Let $\mathcal{J}$ be a $c'_{5.15}\varepsilon_0\rho'_j/10Kk_0$-net of $[0, 1]$ such that

$$|\mathcal{J}| \leq \frac{30Kk_0}{c'_{5.15}\varepsilon_0\rho'_j}.$$

---

[3]See for example [52, (5.7)].

[4]Observe we are applying this lemma when the upper limit is $2L$, according to the definition of S_L, not L. The definition of β is adjusted accordingly below.

As observed earlier, the partition $\{I_k\}_{k \geq 0}$ of the coordinates of v is entirely determined by the sets of indices τ and σ. To approximate all $v \in \mathcal{K}$, we define the preliminary set

$$\mathcal{M}' = \bigcup_{\tau, \sigma \in [n]: |\tau| = m, |\sigma| = m/4} \left\{ x_0 + \sum_{k=1}^{k_0} t_k y_k : x_0 \in \mathcal{N}_0, y_k \in \mathcal{N}_k, t_k \in \mathcal{J} \right\}. \qquad (5.7)$$

We currently have no guarantee that

$$\mathcal{M}' \subset \mathrm{Incomp}(m, 2^{j-1}\rho, 2^j \rho) \cap S_L.$$

However, this is easily fixed. If there exists $x \in S_L$ such that

$$\|x - m\|_2 \leq \frac{c_{5.15}\rho'_j \varepsilon_0}{15K},$$

we replace m by any such x. Otherwise, we discard m. This creates a new net $\mathcal{M}$ such that $|\mathcal{M}| \leq |\mathcal{M}'|$. This completes Step 1.

We now enter Step 2 of the proof and upper bound the size of $\mathcal{M}$. We may combinatorially determine the size of $\mathcal{M}$ using the sizes of the $\mathcal{N}_k$ and $\mathcal{J}$. This leads to the following bound on the cardinality of our net:

$$|\mathcal{M}| \leq \binom{n}{m} \binom{n}{m/4} \left(\frac{30K}{c'_{5.15}\varepsilon_0 \rho'} \right)^{m+\alpha n} \prod_{k=1}^{k_0} \left[\left(12 + \frac{\bar{c}2L}{\sqrt{\alpha n}} \right)^{\alpha n} \log(L) \frac{30Kk_0}{c'_{5.15}\varepsilon_0 \rho'_j} \right]. \qquad (5.8)$$

The combinatorial factors come from the choices of τ and σ in (5.7).

We now proceed to simplify this bound. From the elementary bound

$$\binom{n}{k} \leq \exp(k \log(en/k))$$

we have the following exponential bound for $|\mathcal{M}|$:

$$|\mathcal{M}| \leq \exp\left(2m \log(4en/m) + (m + \alpha n + k_0) \log(30K/c'_{5.15}\varepsilon_0 \rho'_j) \right.$$
$$\left. + k_0 \log(\log(L)) + k_0 \log(k_0) \right) \times \left(12 + \frac{\bar{c}2L}{\sqrt{\alpha n}} \right)^{n-m}.$$

For the second factor, we recalled that $|I_0| \geq m$, so that the product from 1 to k_0 in (5.8) has at most $n - m$ individual terms. Using $L \leq \exp(2\alpha^{-1})$, $m = \alpha n$, $k_0 \leq \alpha n$, and $k_0 \leq \alpha^{-1}$ (from (5.2)), we find

$$|\mathcal{M}| \leq \exp\left(n \left[2\alpha \log(4e/\alpha) + 3\alpha \log(30K/c'_{5.15}\varepsilon_0 \rho'_j) + \frac{1}{n\alpha} \log(2/\alpha^2) \right] \right)$$
$$\times \left(12 + \frac{\bar{c}2L}{\sqrt{\alpha n}} \right)^{n-m}.$$

Recall that ρ'_j was defined in terms of ρ in Lemma 5.12, and $\log(1/\rho) = O(\log n / \log \log n)$ by Remark 4.5. Note also that $\log(1/\alpha) = O(\log n)$. Then there exists $C > 0$ such that

$$2\alpha \log(4e/\alpha) \leq C, \qquad 3\alpha \log(30K/c'_{5.15}\rho'_j) \leq C, \qquad \frac{1}{n\alpha} \log(2/\alpha^2) \leq C.$$

From this, we find

$$|\mathcal{M}| \leq \exp\left(n \left[C + 3\alpha \log(1/\varepsilon_0) \right] \right) \times \left(12 + \frac{\bar{c}2L}{\sqrt{\alpha n}} \right)^{n-\alpha n}. \qquad (5.9)$$

This completes Step 2.

We now begin Step 3 of the outline and prove the result for all the points in our net $\mathcal{M}$. Set

$$P = \mathbb{P}\left(\exists x \in \mathcal{M} \text{ s.t. } \|(M_n - \lambda)x\|_2 \leq c_{5.15}\varepsilon_0\rho_j'\sqrt{pn}\right).$$

By Proposition 5.5 applied with $\varepsilon = c_{5.15}\varepsilon_0$, for any $v \in \mathcal{M}$ and k such that $1 \leq k \leq k_0$,

$$\mathbb{P}\left(\|(M_n - \lambda)v\|_2 \leq c_{5.15}\varepsilon_0\rho_j'\sqrt{pn}\right) \leq C_{5.5}^{n-\lceil \alpha n \rceil}\left(c_{5.15}\varepsilon_0 + \frac{1}{\sqrt{p}D(v_{I_k}/\|v_{I_k}\|_2)}\right)^{n-\lceil \alpha n \rceil},$$

where we recall from Lemma 5.12 that $\rho_j' \leq \|v_{I_k}\|_2$. Since $v \in S_L$, by the definition of S_L we find there exists $1 \leq k \leq k_0$ such that $D(v_{I_k}/\|v_{I_k}\|_2) > L$. We use this k in the above expression to find

$$\mathbb{P}\left(\|(M_n - \lambda)v\|_2 \leq c_{5.15}\varepsilon_0\rho_j'\sqrt{pn}\right) \leq C_{5.5}^{n-\lceil \alpha n \rceil}\left(c_{5.15}\varepsilon_0 + \frac{1}{\sqrt{p}L}\right)^{n-\lceil \alpha n \rceil}.$$

Straightforward computations show

$$\frac{c_{5.15}'\sqrt{\alpha n}}{L} \geq \frac{2C_{5.5}}{\sqrt{p}L} \quad \text{and} \quad \frac{c_{5.15}'\sqrt{\log r}(\log\log n)}{\alpha^2 r} \geq \frac{2C_{5.5}}{\sqrt{p}L}. \tag{5.10}$$

Recall that ε_0 as defined as the minimum of the two upper bounds in (5.10), so

$$\varepsilon_0 \geq \frac{2C_{5.5}}{\sqrt{p}L}. \tag{5.11}$$

Then

$$\mathbb{P}\left(\|(M_n - \lambda)x\|_2 \leq c_{5.15}\varepsilon_0\rho_j'\sqrt{pn}\right) \leq C_{5.5}^{n-\lceil \alpha n \rceil}\left(c_{5.15}\varepsilon_0 + (2C_{5.5})^{-1}\varepsilon_0\right)^{n-\lceil \alpha n \rceil}.$$

Setting $c_{5.15} = (2C_{5.5})^{-1}$ and applying a union bound over all elements $x \in \mathcal{M}$, we obtain

$$P \leq |\mathcal{M}|\varepsilon_0^{n-\alpha n}. \tag{5.12}$$

To bound $|\mathcal{M}|\varepsilon_0^{n-\alpha n}$ from (5.12), we use (5.9) and divide into two cases. First, suppose $\frac{2\bar{c}L}{\sqrt{\alpha n}} \leq 1$. By (5.9), we have

$$|\mathcal{M}| \leq \exp\left(n\left[C + 3\alpha\log(1/\varepsilon_0)\right]\right) \times 13^n. \tag{5.13}$$

Combining this with (5.12) and absorbing the 13^n into the exponential yields

$$P \leq \exp\left(n\left[C + 3\alpha\log(1/\varepsilon_0)\right]\right) \times \varepsilon_0^{n-\alpha n},$$

so

$$P \leq \exp\left(n\left[C + 3\alpha\log(1/\varepsilon_0) - (1-\alpha)\log(1/\varepsilon_0)\right]\right) \leq \exp(-c_{5.15}''n).$$

In the last line we used $\alpha = o(1)$ and $\varepsilon_0 \to 0$ (the latter is by direct calculation), so $\log(1/\varepsilon_0) \to \infty$ and the term inside the brackets tends to $-\infty$.

For the case $\frac{2\bar{c}L}{\sqrt{\alpha n}} > 1$, recalling the definition of ε_0 and that $m = \alpha n$ gives

$$P \leq \exp\left(n\left[C + 3\alpha\log(1/\varepsilon_0)\right]\right) \times \left(\frac{13\bar{c}L}{\sqrt{\alpha n}}\right)^{n-\alpha n}\left(\frac{c_{5.15}'\sqrt{\alpha n}}{L}\right)^{n-\alpha n}, \tag{5.14}$$

$$P \leq \exp\left(n\left[C + 3\alpha\log(1/\varepsilon_0)\right]\right) \times \left(13\bar{c}c_{5.15}'\right)^{n-\alpha n}. \tag{5.15}$$

Now (5.11) shows that

$$\frac{1}{\varepsilon_0} \leq \frac{\sqrt{p}L}{2C_{5.5}} \leq \frac{\exp(\alpha^{-1})}{2C_{5.5}}.$$

This, along with the stipulated range of α, implies that

$$3\alpha \log(1/\varepsilon_0) < C.$$

Therefore, taking $c'_{5.15}$ small enough in (5.15), we have

$$P \leq \exp(-c''_{5.15} n).$$

This completes Step 3.

We now proceed to Step 4. Having shown the result for all the points in the net, we now extend to the entire level set $\mathcal{K}$. Again, we divide into cases.

We assume first that

$$\frac{c'_{5.15}\sqrt{\alpha n}}{L} \leq \frac{c'_{5.15}\sqrt{\log r}(\log\log n)}{\alpha^2 r}, \quad \text{so that} \quad \varepsilon_0 = \frac{c'_{5.15}\sqrt{\alpha n}}{L}. \tag{5.16}$$

For any $w \in \mathcal{K}$, let $m \in \mathcal{M}$ be the closest element of the net $\mathcal{M}$. Then, by the definition of $\mathcal{M}$,

$$
\begin{aligned}
\|w - m\|_2 &\leq \frac{c_{5.15}\rho'_j\varepsilon_0}{10K} + \sum_{k=1}^{k_0}\left(\left\|w_{I_k} - \|w_{I_k}\|_2 y_k\right\|_2 + \left\|\|w_{I_k}\|_2 y_k - t_k y_k\right\|_2\right) \\
&\leq \frac{c_{5.15}\rho'_j\varepsilon_0}{10K} + \sum_{k=1}^{k_0}\left(\left\|\frac{w_{I_k}}{\|w_{I_k}\|_2} - y_k\right\|_2 \|w_{I_k}\|_2 + \left\|\|w_{I_k}\|_2 y_k - t_k y_k\right\|_2\right) \\
&\leq \frac{c_{5.15}\rho'_j\varepsilon_0}{10K} + k_0\rho'_j\alpha^{-1}(c'_{4.4})^{-2}2^3\beta + k_0\frac{c'_{5.15}\rho'_j\varepsilon_0}{10Kk_0} \\
&\leq \frac{c_{5.15}\rho'_j\varepsilon_0}{5K} + C_\gamma(c'_{4.4})^{-2}\frac{\rho'_j}{\alpha^{5/2}L\sqrt{p}} \\
&\leq \frac{c_{5.15}\rho'_j\varepsilon_0}{2K}.
\end{aligned}
$$

In the third inequality, we used that there are k_0 terms in the sum, that the y_k form a β-net, and the upper bound on $\|w_k\|_2$ from (5.4). In the fourth inequality, we used $k_0 \leq \alpha^{-1}$ from (5.2) and the inequality

$$\beta \leq \frac{C_\gamma \alpha^{-1/2}}{L\sqrt{p}}, \tag{5.17}$$

where C_γ is a constant that depends only on γ. The inequality (5.17) follows from the definition of β and the hypothesized upper bound $\log(\sqrt{p}L) \leq \alpha^{-1}$ on L. The last inequality follows by direct calculation using the value of ε_0 given in (5.16) and the assumed lower bound on α.

For the other case, suppose

$$\frac{c'_{5.15}\sqrt{\alpha n}}{L} \geq \frac{c'_{5.15}\sqrt{\log r}(\log\log n)}{\alpha^2 r}, \quad \text{so that} \quad \varepsilon_0 = \frac{c'_{5.15}\sqrt{\log r}(\log\log n)}{\alpha^2 r}. \tag{5.18}$$

For any $w \in \mathcal{K}$, let $m \in \mathcal{M}$ be the closest element of the net $\mathcal{M}$. Then, by the definition

of $\mathcal{M}$,

$$
\begin{aligned}
\|w - m\|_2 &\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{10K} + \sum_{k=1}^{k_0}\left(\left\|w_{I_k} - \|w_{I_k}\|_2 y_k\right\|_2 + \left\|\|w_{I_k}\|_2 y_k - t_k y_k\right\|_2\right) \\
&\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{10K} + \sum_{k=1}^{k_0}\left(\left\|\frac{w_{I_k}}{\|w_{I_k}\|_2} - y_k\right\|_2 \|w_{I_k}\|_2 + \left\|\|w_{I_k}\|_2 y_k - t_k y_k\right\|_2\right) \\
&\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{10K} + k_0\rho_j'\alpha^{-1}(c_{4.4}')^{-2}2^3\beta + k_0\frac{c_{5.15}\rho_j'\varepsilon_0}{10Kk_0} \\
&\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{5K} + k_0\rho_j'\alpha^{-1}(c_{4.4}')^{-2}2^3\frac{\sqrt{\log(4\sqrt{\gamma p L})}}{\sqrt{\gamma p}L} \\
&\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{5K} + \rho_j'\alpha^{-2}(c_{4.4}')^{-2}2^4\frac{\sqrt{\log(4\sqrt{\gamma}r)}}{\sqrt{\gamma}r} \\
&\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{5K} + \left((c_{4.4}')^{-2}(c_{5.15})^{-1}\gamma^{-1/2}2^5\right)\rho_j'\varepsilon_0(\log\log n)^{-1} \\
&\leq \frac{c_{5.15}\rho_j'\varepsilon_0}{2K}.
\end{aligned}
$$

In the third line, we used that there are k_0 terms in the sum, that the y_k's form a β-net, and the upper bound on $\|w_k\|_2$ from (5.4). The fourth line follows from the definition of β. The fifth line is a result of the observation that $\sqrt{\log x}/x$ is a decreasing function for large x, $r \to \infty$, and $r < L\sqrt{p}$. We also used the bound $k_0 \leq \alpha^{-1}$ from (5.2). In the sixth line, we used the definition of ε_0 in (5.18). For the last line, we used $(\log\log n)^{-1} = o(1)$ and took n large enough.

Therefore, if $\|(M_n - \lambda)w\|_2 \geq 2c_{5.15}\varepsilon_0\sqrt{pn}$, then using Lemma 4.3,

$$
\|(M_n - \lambda)m\|_2 \geq 2c_{5.15}\varepsilon_0\sqrt{pn} - \|M_n - \lambda\|\frac{c_{5.15}\varepsilon_0}{2K} \geq c_{5.15}\varepsilon_0\sqrt{pn},
$$

with exponentially small error probability, which contradicts the conclusion of Step 3 above. After adjusting $c_{5.15}$ by a factor of 2, this completes the proof. $\qquad\square$

Remark 5.16. As noted in Remark 2.5, the optimal result should permit p as small as $C\log(n)/n$. The restriction that $p \geq C\log^{7+\nu} n/n$ in the above proof comes from the requirement that $\varepsilon_0 \to 0$.

We now extend the previous result to all vectors with small LCD.

Proposition 5.17. *Fix $\nu > 0$. There exists a constant $c_{5.17} > 0$ such that for $p \geq C_{5.15}\frac{\log^{7+\nu} n}{n}$, $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$, $j \in \mathbb{N}$ and for any*

$$
(np)^{-1/(7+\nu)} \leq \alpha \leq \frac{c_{5.15}'}{\log n}
$$

the following holds. The probability that there exists $v \in \mathrm{Incomp}(\alpha n, \rho)$ such that

$$
\|(M_n - \lambda)v\|_2 \leq c_{5.15}\varepsilon_1\rho_1'\sqrt{pn} \text{ and } \widehat{D}(v) \leq p^{-1/2}\exp(\alpha^{-1})
$$

is at most $\exp(-c_{5.17}n)$ for $n \geq (c_{5.17})^{-1}$, where

$$
\varepsilon_1 = \exp(-c_{5.17}n^{1/7}).
$$

Proof. We set $D_0 = c_{4.4}'2^{-5}\alpha^{3/2}n^{1/2}$ and recall that $\widehat{D}(v) \geq D_0$ by (5.6). We can decompose the relevant vectors as

$$
\bigcup_{j'=0}^{\log_2\left(p^{-1/2}\exp(\alpha^{-1})\right)}\bigcup_{j=0}^{\log_2\rho^{-1}}\left(\mathrm{Incomp}(m, 2^j\rho, 2^{j+1}\rho) \cap S_{2^{j'}D_0}\right),
$$

where we used $D_0 \geq 1$. Recall $\log(1/\rho) = O(\log n / \log \log n)$ by Remark 4.5. Similarly, the number of j' indices in the union is $O(\log n)$ because each of $\log_2 p^{-1/2}$ and $\log_2 \exp(\alpha^{-1})$ are $O(\log n)$. Therefore, taking a union bound, applying Proposition 5.15, and observing $\rho'_j \geq \rho'_1$ and $\varepsilon_0(L) \geq \varepsilon_1$ for the $\varepsilon_0(L)$ defined in Proposition 5.15 yields the result. $\qquad\square$

5.4 Eigenvector bound

We now come to a key proposition used in the proof of the main theorem.

Proposition 5.18. *For M_n as in Definition 2.1, there exists a constant $c_{5.18} > 0$ such that for*

$$(np)^{-1/(7+\nu)} \leq \alpha \leq \frac{c'_{5.15}}{\log n},$$

the probability that M_n has an eigenvector v such that

$$v \notin \mathrm{Comp}(\alpha n, \rho) \cup \mathrm{Dom}(\alpha n, c_{4.4}) \text{ and } \widehat{D}(v) \leq p^{-1/2} \exp(\alpha^{-1})$$

is at most $\exp(-c_{5.18}n)$, for $n \geq (c_{5.18})^{-1}$.

Proof. Consider a $c_{5.15}\varepsilon_1\rho'_1\sqrt{pn}$-net of $[-K\sqrt{pn}, K\sqrt{pn}]$, where ε_1 was defined in Proposition 5.17. For an eigenvalue $\lambda \in [-K\sqrt{pn}, K\sqrt{pn}]$, there exists a point of the net λ_0 such that for corresponding eigenvector v we have

$$\|(M_n - \lambda_0)v\|_2 = |\lambda - \lambda_0| \leq c_{5.15}\varepsilon_1\rho'_1\sqrt{pn}.$$

However, by a union bound and Proposition 5.17, the probability of this event is bounded by $\exp(-c_{5.18}n)$ for some $c_{5.18} > 0$. By Lemma 4.3, decreasing the value of $c_{5.18}$ can account for the event that there exists an eigenvalue of M_n outside the interval $[-K\sqrt{pn}, K\sqrt{pn}]$. This concludes the proof. $\qquad\square$

6 Proofs of Main Results

6.1 Proof of Theorem 2.2

In preparation for the main proof, we record the following lemma from [43].

Lemma 6.1 ([43, Lemma 6.1]). *For any $v \in \mathrm{Incomp}(m, \rho)$,*

$$\left| \left\{ i : \frac{\rho^2}{\sqrt{2n}} \leq |v_i| \leq \frac{1}{\sqrt{m}} \right\} \right| \geq \frac{m\rho^2}{2}.$$

Proof of Theorem 2.2. We repeat the decomposition described in Section 3. Let

$$M_n = \begin{pmatrix} M_{n-1} & X \\ X^T & m_{nn} \end{pmatrix}, \tag{6.1}$$

where $X = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$. Let $v = (x, a)$ (where $x \in \mathbb{R}^{n-1}$ and $a \in \mathbb{R}$) be the unit eigenvector associated to $\lambda_i(M_n)$. Because v is an eigenvector with eigenvalue λ_i,

$$\begin{pmatrix} M_{n-1} & X \\ X^T & m_{nn} \end{pmatrix} \begin{pmatrix} x \\ a \end{pmatrix} = \lambda_i(M_n) \begin{pmatrix} x \\ a \end{pmatrix}.$$

Considering the top $n-1$ coordinates gives

$$(M_{n-1} - \lambda_i(M_n))x + aX = 0.$$

Let w be the eigenvector of M_{n-1} corresponding to $\lambda_i(M_{n-1})$. After multiplying on the left by w^T, we arrive at

$$|aw^TX| = |w^T(M_{n-1} - \lambda_i(M_n))x| = |\lambda_i(M_{n-1}) - \lambda_i(M_n)||w^Tx|. \tag{6.2}$$

Since $|w^T x| \leq 1$ by the Cauchy–Schwarz inequality, this implies

$$|w^T X| \leq \frac{1}{|a|} |\lambda_i(M_{n-1}) - \lambda_i(M_n)|. \tag{6.3}$$

By the Cauchy interlacing law, we must have $\lambda_i(M_n) \leq \lambda_i(M_{n-1}) \leq \lambda_{i-1}(M_n)$. For any $\hat{\delta} > 0$, let $\mathcal{E}_i = \mathcal{E}_i\left(\hat{\delta}\right)$ denote the event that

$$\lambda_{i+1} - \lambda_i \leq \hat{\delta}\sqrt{\frac{p}{n}}. \tag{6.4}$$

On $\mathcal{E}_i$, (6.3) implies

$$|w^T X| \leq \hat{\delta}\sqrt{\frac{p}{n}} \frac{1}{|a|}. \tag{6.5}$$

Now note that the decomposition (6.1) can be done along any coordinate, not just the last. For any $A > 0$, let n_A be the number of coordinates with absolute value at least A, and let N be a parameter. Therefore, repeating the argument leading to (6.5) with the coordinate a chosen uniformly at random, and considering the probability that we choose a coordinate with absolute value at least A, and $\mathcal{E}_i$ obtains, we find

$$\mathbb{P}(\mathcal{E}_i) = \mathbb{P}(\mathcal{E}_i \cap \{n_A \geq N\}) + \mathbb{P}(\mathcal{E}_i \cap \{n_A < N\}) \tag{6.6}$$

$$\leq \frac{n}{N}\mathbb{P}\left(|w^T X| \leq \hat{\delta}\sqrt{\frac{p}{n}} \frac{1}{A}\right) + \mathbb{P}(n_A < N). \tag{6.7}$$

Setting $m = c_{4.4}n$ in Proposition 4.6 shows that any eigenvector v will not be in $\mathrm{Comp}(c_{4.4}n, \rho)$ with exponentially high probability. When $v \notin \mathrm{Comp}(c_{4.4}n, \rho)$, by Lemma 6.1, there are greater than $c_{4.4}n\rho^2/2$ coordinates whose absolute values are larger than $\rho/\sqrt{2n}$. We set $N = c_{4.4}n\rho^2/2$ and $A = \rho/\sqrt{2n}$ in (6.7) to find

$$\mathbb{P}(\mathcal{E}_i) \leq \frac{2}{c_{4.4}\rho^2}\mathbb{P}\left(|w^T X| \leq \hat{\delta}\rho^{-1}\sqrt{2p}\right) + \exp(-c_{4.6}pn). \tag{6.8}$$

With probability at least $1 - \exp(-c_{5.18}pn)$,

$$\widehat{D}(w) \geq p^{-1/2}\exp(\alpha^{-1})$$

by Proposition 4.6 (applied with $m = \alpha n$) and Proposition 5.18. At this point, we would like to apply Proposition 5.4 to control the probability $\mathbb{P}\left(|w^T X| \leq \hat{\delta}\rho^{-1}\sqrt{2p}\right)$ in (6.8). However, this proposition applies to the LCD $D(w)$, not the regularized LCD $\widehat{D}(w)$, so a slightly more delicate argument is required.

By the definition of regularized LCD, there exists some subset J of coordinate indices such that

$$D\left(\frac{w_J}{\|w_J\|_2}\right) \geq p^{-1/2}\exp(\alpha^{-1}).$$

To adjust for the regularized LCD, we observe that conditioning on a subset of X can only increase the Lévy function $\mathcal{L}(w^T X, \varepsilon)$ for any $\varepsilon > 0$. We condition on all the random variables in X whose indices do not lie in the subset J. Also, to apply Proposition 5.4, we need to normalize this subset to be on the unit sphere. Therefore, by Proposition 5.4,

$$\mathcal{L}(w^T X, \hat{\delta}\rho^{-1}\sqrt{2p}) \leq \mathcal{L}\left(\frac{w_J^T}{\|w_J\|_2}X_J, \frac{\hat{\delta}\rho^{-1}\sqrt{2p}}{\|w_J\|_2}\right) \leq \frac{2\sqrt{2}C_{5.4}\hat{\delta}\rho^{-1}}{\|w_J\|_2}, \tag{6.9}$$

for all $\hat{\delta} \geq \rho e^{-\alpha^{-1}}/\sqrt{2}$. By Lemma 5.12, $\|w_J\|_2 \geq c'_{4.4}2^{-3}\rho\alpha$. Therefore, putting (6.9) into (6.8), we find

$$\mathbb{P}(\mathcal{E}_i) \leq \frac{32\sqrt{2}}{c_{4.4}c'_{4.4}\rho^4\alpha}C_{5.4}\hat{\delta} + \exp(-c_{4.6}pn). \tag{6.10}$$

We set $\delta = \hat{\delta}\rho^{-4}$. Then the above holds for $\delta \geq \rho^{-3}e^{-\alpha^{-1}}/\sqrt{2}$. Recall that $\rho^{-3} = \exp(O(\log n/\log\log n))$. Thus, we obtain the theorem after lowering $c'_{2.2}$, which constrains the range of α. $\qquad\square$

6.2 Proof of Theorem 2.6

Let $G(n,p)$ denote the Erdős–Rényi random graph on n vertices with edge probability p, and let A_n denote the adjacency matrix of $G(n,p)$. In other words, A_n is a symmetric matrix of Bernoulli variables with parameter p, with all 0 entries on the diagonal. We have $\mathbb{E}A_n = p(J_n - I_n)$ where J_n is the matrix of all ones, so our main theorem does not apply. However, only small modifications are necessary to handle this case, which we detail in this section, following closely the analogous argument in [43, Section 8].

First, we observe that Proposition 4.4 can be adapted so that the proposition holds for A_n in place of M_n. This was proved in [43, Appendix B]. It follows that Proposition 4.6 also holds for A_n (by repeating the proof of Proposition 4.6 using the analogue of Proposition 4.4 for A_n).

Next, we claim that Proposition 5.17 can be adapted to hold for the matrix $A_n - p(J_n - I_n)$ in place of M_n, *with the additional restriction that we must suppose $p \leq 1/2$*. The restriction is due to the fact that we will write the off-diagonal entries of this matrix as $a_{ij} = \delta_{ij}\xi_{ij}$, where δ is Bernoulli with parameter $2p$ and ξ_{ij} is Bernoulli with parameter $1/2$ (as in the definition of M_n). Our arguments for Proposition 5.17 revolved around Lévy concentration and nets. The use of Lévy concentration in Proposition 5.5 does not need to be modified for the random graph case, since it is invariant under changes in the mean of the matrix.[5] For the nets, we required the operator norm bound Lemma (4.3); we claim the analogue of this statement for $A_n - p(J_n - I_n)$ also holds. A straightforward modification of the proof of [9, Theorem 1.7] shows

$$\mathbb{P}(\|A_n - p(J_n - I_n)\|_2 \geq K'\sqrt{pn}) \leq \exp(-c'pn) \tag{6.11}$$

for some $K', c' > 0$. We obtain that Proposition 5.17 holds for $A_n - p(J_n - I_n)$, if $p \leq 1/2$.

Additionally, we need a slight generalization of Proposition 5.17, which lower bounds not just $\|(A_n - p(J_n - I_n) - \lambda)v\|_2$, but

$$\|(A_n - p(J_n - I_n) - \lambda)v - x\|_2 \tag{6.12}$$

for any fixed vector x. This generalization holds because the high probability lower bounds used to prove Proposition 5.17 come from Proposition 5.5, and the latter proposition concerns Lévy concentration, which is by definition translation invariant.

We now turn to the proof of Theorem 2.6.

Proof of Theorem 2.6. Above, we established that the analogue of Proposition 5.17 holds for $A_n - p(J_n - I_n)$, if $p \leq 1/2$. This restriction motivates the following division into cases.

Case I: $p \leq 1/2$. Our preliminary goal to is establish that Proposition 5.18 holds for A_n. We have

$$\{J_n x : x \in \mathbb{S}^{n-1}\} = \{\theta \cdot \mathbb{1} : \theta \in [-n, n]\} \tag{6.13}$$

where $\mathbb{1}$ is the vector $(1, \ldots, 1)$ of all ones. Set $\mathcal{X}_n = \{\kappa \cdot \mathbb{1} : \kappa \in [-pn, pn]\}$. Let $\mathcal{B}$ be a $c_{5.17}\varepsilon_0\rho'\sqrt{pn}$-net of $\mathcal{X}_n$ such that

$$|\mathcal{B}| \leq \frac{4pn}{c_{5.17}\varepsilon_0\rho'\sqrt{pn}} \leq C\exp(cn^{1/7}). \tag{6.14}$$

---

[5]However, it does require the aforementioned decomposition $a_{ij} = \delta_{ij}\xi_{ij}$, giving rise to the $p \leq 1/2$ restriction.

For $x, x' \in \mathcal{X}_n$, the reverse triangle inequality yields

$$\left| \|(A_n - p(J_n - I_n) - \lambda)v - x\|_2 - \|(A_n - p(J_n - I_n) - \lambda)v - x'\|_2 \right| \leq \|x - x'\|_2, \quad (6.15)$$

so any $(A_n - p(J_n - I_n) - \lambda)v - y$ with $y \in \mathcal{X}_n$ can be well approximated by $(A_n - p(J_n - I_n) - \lambda)v - x$ for some $x \in \mathcal{B}$.

Define

$$S_D = \left\{ v \in \mathrm{Incomp}(\alpha n, \rho) : \widehat{D}(v) \leq p^{-1/2} \exp(\alpha^{-1}) \right\}.$$

By (6.15), a union bound over the net $\mathcal{B}$, and the analogue of Proposition 5.17 for (6.12) stated above, we obtain

$$\mathbb{P}\left(\inf_{x \in \mathcal{X}_n} \inf_{v \in S_D} \|(A_n - p(J_n - I_n) - \lambda)v - x\|_2 \leq c_{5.17}\varepsilon_0\rho_1'\sqrt{pn} \right) \leq \exp(-cn) \quad (6.16)$$

for any single $\lambda \in [-K'\sqrt{pn}, K'\sqrt{pn}]$. After observing that

$$\inf_{x \in \mathcal{X}_n} \inf_{v \in S_D} \|(A_n - p(J_n - I_n) - \lambda)v - x\|_2 \leq \inf_{v \in S_D} \|(A_n - (\lambda - p))v\|_2, \quad (6.17)$$

we find

$$\mathbb{P}\left(\inf_{v \in S_D} \|(A_n - (\lambda - p))v\|_2 \leq c_{5.17}\varepsilon_0\rho_1'\sqrt{pn} \right) \leq \exp(-cn). \quad (6.18)$$

Using (6.18) in place of Proposition 5.17 in the proof of Proposition 5.18, we find that Proposition 5.18 holds for A_n in place of M_n.

We can now repeat the proof of Theorem 2.2 to prove theorem in this case, with the appropriate analogues for A_n substituting for Proposition 5.18 and Proposition 4.6. (The latter was noted at the beginning of Section 6.2.)

Case II: $p > 1/2$. Observe that the adjacency matrix $A_n(p)$ of $G(n, p)$ is equal in distribution to $J_n - I_n - A_n(1 - p)$. Hence controlling

$$\|(A_n(p) - p(J_n - I_n) - \lambda)v\|_2$$

is equivalent to controlling

$$\|(A_n(1 - p) - (1 - p)(J_n - I_n) + \lambda)v\|_2.$$

This reduces the problem to Case I and completes the proof. $\qquad\square$

Remark 6.2. The size of the one-dimensional net $\mathcal{B}$ in (6.14) is compensated by the $\exp(-cn)$ error probability used for the union bound in (6.16). For general finite-rank perturbations by a finite linear combination of matrices of the form $n \cdot vv^T$ for $v \in \mathbb{S}^n$, one simply adds more one-dimensional nets and completes the argument in the same way. However, for perturbations whose rank grows even moderately quickly, the combined size of the necessary supplemental nets becomes too large.

6.3 Proof of Theorem 2.8

The following is essentially Lemma 9.1 of [48]. We provide the proof for completeness.

Lemma 6.3. *For any $A > 0$ there exists $B = B(A) > 0$ such that the following holds with probably at least $1 - O(n^{-A})$. If there exist $\lambda \in \mathbb{R}$ and $v \in \mathbb{S}^{n-1}$ such that $\|(A_n - \lambda)v\| \leq n^{-B}$, then A_n has an eigenvector $u_{i_0} \in \mathbb{S}^{n-1}$ and corresponding eigenvalue λ_{i_0} such that*

$$|\lambda_{i_0} - \lambda| < n^{-B/4} \quad and \quad \|v - u_{i_0}\| < n^{-B/4}.$$

Proof. From our main result, Theorem 2.6, we may suppose that all eigenvalue gaps satisfy $|\lambda_j - \lambda_i| \geq n^{-B/2}$. Let $v = \sum c_i u_i$ express v as a linear combination of unit eigenvectors of A. There must exist i_0 such that $c_{i_0} \geq n^{-1/2}$. So

$$\|(A_n - \lambda)v\| = \left(\sum_{i=1}^{n} c_i^2 (\lambda_i - \lambda)^2 \right)^{1/2} \tag{6.19}$$

implies, assuming $\|(A_n - \lambda)v\| \leq n^{-B}$, that $|\lambda - \lambda_{i_0}| \leq n^{-B+1/2}$. This implies the first conclusion. Then because all gaps satisfy $|\lambda_j - \lambda_i| \geq n^{-B/2}$ we have that $|\lambda - \lambda_i| \geq n^{-B/2}/2$ for all $i \neq i_0$. But then we must have $|c_i| = O(n^{-B/2})$ for $i \neq i_0$, implying the second conclusion. $\qquad\square$

Proof of Theorem 2.8. We follow the proof of Theorem 3.3 in [48]. After adjusting C by adding 1, it suffices to prove the claim for a single coordinate and use a union bound. Write $A = A_n$ and let its first column be (a_{11}, X) where X is a vector of $n-1$ coordinates. Let $v = (v_1, v')$ be an eigenvector with eigenvalue λ so that

$$v_1 m_{11} + (v')^T X = \lambda v_1, \quad (A_{n-1} - \lambda)v' = -v_1 X. \tag{6.20}$$

Suppose that $|v_1| \leq n^{-D}$ where D will be chosen later. By taking D large enough, using that the entries of A are bounded, and adding $O(N^{-D})$ mass to the first component of v' to make it unit norm, it suffices to show that

$$\|(A_{n-1} - \lambda)v'\| \leq n^{-D/2} \quad \text{and} \quad |(v')^T X| \leq n^{-D/2} \tag{6.21}$$

occur jointly with low probability. By Lemma 6.3, if the first condition holds then there exists an eigenvector u' of A_{n-1} with $\|u' - v'\|_2 \leq n^{-D/8}$. Then $|(v')^T X| \leq N^{-D/2}$ implies $|(u')^T X| \leq n^{-D/16}$. We claim this contradicts a statement established in the proof of Theorem 2.2.

In (6.9) and the following lines, we showed

$$\mathbb{P}\left(|w^T X| \leq \delta \rho^3 \sqrt{p} \right) \leq C \frac{\rho^2}{\alpha} \delta, \tag{6.22}$$

where δ was defined below (6.10) (in terms of $\hat{\delta}$). Now we take $\delta = n^{-D/16}/\rho^3 \sqrt{p}$, $\alpha = (np)^{-1/(7+\nu)}$ and $p > C \log^{7+\nu}(n)/n$, which proves the theorem after taking D large enough. $\qquad\square$

References

[1] Yonathan Aflalo, Alex Bronstein, and Ron Kimmel. Graph matching: relax or not? *arXiv preprint* arXiv:1401.7623, 2014.

[2] Amol Aggarwal. Bulk universality for generalized Wigner matrices with few moments. *Probability Theory and Related Fields*, 173(1-2):375–432, 2019. MR3916110

[3] Amol Aggarwal, Patrick Lopatto, and Horng-Tzer Yau. GOE statistics for Lévy matrices. *arXiv preprint* arXiv:1806.07363, 2018. MR4310816

[4] Sanjeev Arora and Aditya Bhaskara. Eigenvectors of random graphs: delocalization and nodal domains. https://theory.epfl.ch/bhaskara/files/deloc.pdf, 2011.

[5] Enrico Au-Yeung. Sparse signal recovery using a new class of random matrices. *Adv. Pure Appl. Math.*, 8(2):79–89, 2017. MR3630095

[6] László Babai, D. Yu. Grigoryev, and David Mount. Isomorphism of graphs with bounded eigenvalue multiplicity. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, pages 310–324. ACM, 1982.

[7] Bubacarr Bah and Jared Tanner. On construction and analysis of sparse random matrices and expander graphs with applications to compressed sensing. *arXiv preprint* arXiv:1307.6477, 2013. MR3124656

[8] Grey Ballard, Aydin Buluc, James Demmel, Laura Grigori, Benjamin Lipshitz, Oded Schwartz, and Sivan Toledo. Communication optimal parallel multiplication of sparse random matrices. In *Proceedings of the twenty-fifth annual ACM symposium on Parallelism in algorithms and architectures*, pages 222–231. ACM, 2013.

[9] Anirban Basak and Mark Rudelson. Invertibility of sparse non-Hermitian matrices. *Adv. Math.*, 310:426–483, 2017. MR3620692

[10] Anirban Basak and Mark Rudelson. Sharp transition of the invertibility of the adjacency matrices of sparse random graphs. *arXiv preprint* arXiv:1809.08454, 2018. MR4265022

[11] F. Bekerman, A. Figalli, and A. Guionnet. Transport maps for β-matrix models and universality. *Comm. Math. Phys.*, 338(2):589–619, 2015. MR3351052

[12] Gérard Ben Arous and Paul Bourgade. Extreme gaps between eigenvalues of random matrices. *Ann. Probab.*, 41(4):2648–2681, 2013. MR3112927

[13] Lucas Benigni and Patrick Lopatto. Optimal delocalization for generalized Wigner matrices. *arXiv preprint* arXiv:2007.09585, 2020.

[14] Paul Bourgade. Extreme gaps between eigenvalues of Wigner matrices. *preprint* arXiv:1812.10376, 2018.

[15] Paul Bourgade, Jiaoyang Huang, and Horng-Tzer Yau. Eigenvector statistics of sparse random matrices. *Electronic Journal of Probability*, 22, 2017. MR3690289

[16] Ziliang Che. Universality of random matrices with correlated entries. *Electronic Journal of Probability*, 22, 2017. MR3629874

[17] Ziliang Che and Benjamin Landon. Local spectral statistics of the addition of random matrices. *Probability Theory and Related Fields*, 175(1-2):579–654, 2019. MR4009717

[18] Ziliang Che and Patrick Lopatto. Universality of the least singular value for sparse random matrices. *Electronic Journal of Probability*, 24, 2019. MR3916329

[19] Ziliang Che and Patrick Lopatto. Universality of the least singular value for the sum of random matrices. *arXiv preprint* arXiv:1908.04060, 2019. MR4242226

[20] Shiu-Yuen Cheng. Eigenfunctions and nodal sets. *Commentarii Mathematici Helvetici*, 51(1):43–55, 1976. MR0397805

[21] Kenneth L. Clarkson and David P. Woodruff. Low-rank approximation and regression in input sparsity time. *J. ACM*, 63(6):Art. 54, 45, 2017. MR3614862

[22] Anirban Dasgupta, Ravi Kumar, and Tamás Sarlós. A sparse Johnson–Lindenstrauss transform. In *STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing*, pages 341–350. ACM, New York, 2010. MR2743282

[23] E. Brian Davies, Josef Leydold, and Peter F. Stadler. Discrete nodal domain theorems. *arXiv preprint* arXiv:math/0009120, 2000. MR1855391

[24] Yael Dekel, James R. Lee, and Nathan Linial. Eigenvectors of random graphs: nodal domains. *Random Structures Algorithms*, 39(1):39–58, 2011. MR2839984

[25] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of Erdős-Rényi graphs II: eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.*, 314(3):587–640, 2012. MR2964770

[26] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of Erdős-Rényi graphs I: local semicircle law. *Ann. Probab.*, 41(3B):2279–2375, 2013. MR3098073

[27] L. Erdős, S. Péché, J. A. Ramirez, and B. Schlein. Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.*, 63(7):895–925, 2010. MR2662426

[28] L. Erdős, B. Schlein, and H.-T. Yau. Unversality of random matrices and local relaxation flow. *Invent. Math.*, 185(1):75–119, 2011. MR2810797

[29] László Erdős, Benjamin Schlein, and Horng-Tzer Yau. Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Not. IMRN*, (3):436–479, 2010. MR2587574

[30] László Erdős and Horng-Tzer Yau. Gap universality of generalized Wigner and β-ensembles. *J. Eur. Math. Soc. (JEMS)*, 17(8):1927–2036, 2015. MR3372074

[31] Renjie Feng, Gang Tian, and Dongyi Wei. Small gaps of GOE. *Geometric and Functional Analysis*, 29(6):1794–1827, 2019. MR4034920

[32] Renjie Feng and Dongyi Wei. Large gaps of CUE and GUE. *arXiv preprint* arXiv:1807.02149, 2018.

[33] Renjie Feng and Dongyi Wei. Small gaps of circular β-ensemble. *arXiv preprint* arXiv:1806.01555, 2018. MR4255136

[34] Alessio Figalli and Alice Guionnet. Universality in several-matrix models via approximate transport maps. *Acta Math.*, 217(1):81–176, 2016. MR3646880

[35] Bernard Helffer, Thomas Hoffmann-Ostenhof, and Susanna Terracini. Nodal domains and spectral minimal partitions. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 26(1):101–138, 2009. MR2483815

[36] Han Huang and Mark Rudelson. Size of nodal domains of the eigenvectors of a $G(n,p)$ graph. *Random Structures & Algorithms*, 2020. MR4129726

[37] Jiaoyang Huang, Benjamin Landon, and Horng-Tzer Yau. Bulk universality of sparse random matrices. *J. Math. Phys.*, 56(12):123301, 2015. MR3429490

[38] Benjamin Landon, Patrick Lopatto, and Jake Marcinek. Comparison theorem for some extremal eigenvalue statistics. *Annals of Probability*, 48(6):2894–2919, 2020. MR4164456

[39] Ji Oon Lee and Kevin Schnelli. Local law and Tracy–Widom limit for sparse random matrices. *Probab. Theory Related Fields*, 171(1-2):543–616, 2018. MR3800840

[40] Chang-Shou Lin. On the second eigenfunctions of the Laplacian in $\mathbb{R}^2$. *Communications in Mathematical Physics*, 111(2):161–166, 1987. MR0899848

[41] Alexander E Litvak, Alain Pajor, Mark Rudelson, and Nicole Tomczak-Jaegermann. Smallest singular value of random matrices and geometry of random polytopes. *Advances in Mathematics*, 195(2):491–523, 2005. MR2146352

[42] Alexander E. Litvak and Omar Rivasplata. Smallest singular value of sparse random matrices. *Studia Math.*, 212(3):195–218, 2012. MR3009072

[43] Kyle Luh and Van Vu. Sparse random matrices have simple spectrum. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 56(4):2307–2328, 2020. MR4164838

[44] Vince Lyzinski, Donniell Fishkind, Marcelo Fiori, Joshua Vogelstein, Carey Priebe, and Guillermo Sapiro. Graph matching: Relax at your own risk. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 38(1), 2016.

[45] Pradipta Mitra. Entrywise bounds for eigenvectors of random graphs. *The Electronic Journal of Combinatorics*, 16(1):131, 2009. MR2558268

[46] E. Müller-Pfeiffer. On the number of nodal domains for eigenfunctions of elliptic differential operators. *Journal of the London Mathematical Society*, 2(1):91–100, 1985. MR0810566

[47] Jelani Nelson and Huy L. Nguyễn. OSNAP: faster numerical linear algebra algorithms via sparser subspace embeddings. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science—FOCS 2013*, pages 117–126. IEEE Computer Soc., Los Alamitos, CA, 2013. MR3246213

[48] Hoi Nguyen, Terence Tao, and Van Vu. Random matrices: tail bounds for gaps between eigenvalues. *Probab. Theory Related Fields*, 167(3-4):777–816, 2017. MR3627428

[49] Hoi H. Nguyen and Van H. Vu. Small ball probability, inverse theorems, and applications. In *Erdős centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 409–463. János Bolyai Math. Soc., Budapest, 2013. MR3203607

[50] Sean O'Rourke, Van Vu, and Ke Wang. Eigenvectors of random matrices: a survey. *Journal of Combinatorial Theory, Series A*, 144:361–442, 2016. MR3534074

[51] Natesh S Pillai and Jun Yin. Universality of covariance matrices. *The Annals of Applied Probability*, 24(3):935–1001, 2014. MR3199978

[52] Gilles Pisier. *The volume of convex bodies and Banach space geometry*, volume 94. Cambridge University Press, 1999. MR1036275

[53] Mark Rudelson. Invertibility of random matrices: norm of the inverse. *Ann. of Math. (2)*, 168(2):575–600, 2008. MR2434885

[54] Mark Rudelson. Delocalization of eigenvectors of random matrices. lecture notes. *arXiv preprint* arXiv:1707.08461, 2017. MR3971158

[55] Mark Rudelson and Konstantin Tikhomirov. The sparse circular law under minimal assumptions. *Geometric and Functional Analysis*, 29(2):561–637, 2019. MR3945840

[56] Mark Rudelson and Roman Vershynin. The Littlewood–Offord problem and invertibility of random matrices. *Advances in Mathematics*, 218(2):600–633, 2008. MR2407948

[57] M. Shcherbina. Change of variables as a method to study general β-models: bulk universality. *J. Math. Phys.*, 55(4):043504, 23, 2014. MR3390602

[58] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, 298(2):549–572, 2010. MR2669449

[59] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011. MR2784665

[60] Terence Tao. The asymptotic distribution of a single eigenvalue gap of a Wigner matrix. *Probab. Theory Related Fields*, 157(1-2):81–106, 2013. MR3101841

[61] Terence Tao and Van Vu. Random covariance matrices: universality of local statistics of eigenvalues. *Ann. Probab.*, 40(3):1285–1315, 2012. MR2962092

[62] Terence Tao and Van Vu. Random matrices: the four-moment theorem for Wigner ensembles. In *Random matrix theory, interacting particle systems, and integrable systems*, volume 65 of *Math. Sci. Res. Inst. Publ.*, pages 509–528. Cambridge Univ. Press, New York, 2014. MR3380700

[63] Terence Tao and Van Vu. Random matrices have simple spectrum. *Combinatorica*, 37(3):539–553, 2017. MR3666791

[64] Terence Tao and Van H. Vu. Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)*, 169(2):595–632, 2009. MR2480613

[65] Shinji Umeyama. An eigendecomposition approach to weighted graph matching problems. *IEEE transactions on pattern analysis and machine intelligence*, 10(5):695–703, 1988.

[66] Roman Vershynin. Invertibility of symmetric random matrices. *Random Structures Algorithms*, 44(2):135–182, 2014. MR3158627

[67] Feng Wei. Investigate invertibility of sparse symmetric matrix. *arXiv preprint* arXiv:1712.04341, 2017.

[68] Philip Matchett Wood. Universality and the circular law for sparse random matrices. *Ann. Appl. Probab.*, 22(3):1266–1300, 2012. MR2977992

[69] Jun Yin, L Erdős, and H.-T. Yau. Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math*, 229(3):1435–1515, 2012. MR2871147